APPENDIX A
"CLEAN" VERSION OF EACH PARAGRAPH/SECTION/CLAIM
37 C.F.R. § 1.121(b)(ii) AND (c)(i)

SPECIFICATION:

Paragraph at page 1, line 5 to page 2, line 18:

The present invention relates to a method of reproducing audio signals or audio/video signals and a reproducing apparatus for the same, and more particularly to a method of processing audio signals capable of reproducing the audio signals without causing noticeable tone variation during the reproducing of the audio signals or the audio/video signals at a high or low speed that is different than the normal playback speed.

BACKGROUND OF THE INVENTION

Video and audio program signals are converted to a digital format, compressed, encoded and multiplexed in accordance with an established algorithm or methodology. The compressed digital system signal, i.e., bitstream, includes a video portion, an audio portion, and an informational portion. Such data is transmitted to a reproducing apparatus via a transmission line or by being stored in a recording medium. A digital reproducing apparatus such as a digital versatile disc (DVD) system, a digital video cassette recorder (VCR) or a computer system incorporated with a multimedia player solution for reproducing multimedia data obtained by multiplexing video data and audio data is provided with a decoding means for reproducing the aforementioned bitstream. This decoding means demultiplexes, de-compresses and decodes the bitstream in accordance with the compression algorithm to supply it as a reproducible signal. The decoded video and audio signals are outputted to a reproducing apparatus such as a screen or a speaker for presentation to the user.

The compressing and encoding of the video and audio signals are performed by a suitable encoder which implements a selected data compression algorithm that conforms to a recognized standard or specification agreed to among the senders and receivers of digital video data. Highly efficient compression standards have been developed by the Moving Pictures Experts Group (MPEG), including MPEG-1 and MPEG-2, which have been continuously improved to suggest MPEG-4. The MPEG standards enable the high speed or low speed reproduction forward or

backward in addition to the normal playback mode in the VCR, DVD or similar multimedia recording/reproducing apparatus.

The MPEG standards define a proposed synchronization scheme based on an idealized decoder known as a standard target decoder (STD). Video and audio data units or frames are referred to as access units (AU) in encoded form, and as presentation units (PU) in unencoded or decoded form. In the idealized decoder, video and audio data presentation units are taken from elementary stream buffers and instantly presented at the appropriate presentation time to the user. A presentation time stamp (PTS) indicating the proper presentation time of a presentation unit is transmitted in an MPEG packet header as a part of the system syntax.

Paragraph at page 5, line 4 to page 6, line 4:

The tone variation arises because the conventional reproducing system of fast or slow reproduction mode simply extends or compresses the presentation time interval of respective audio signals in the time scale. What's worse, any other signal processing is separately applied for preventing the tone variation. In other words, an additional scheme is further required for preventing the tone variation during the fast or slow reproduction mode.

SUMMARY OF THE INVENTION

In considering the above-enumerated problems of the prior art, an object of the present invention is to provide a reproducing method using a filtering processing of audio data capable of reproducing an audio signal or an audio signal incorporated with a moving picture, in case of varying a playback speed into the fast or slow mode, in a tone substantially identical with that of a normal playback mode, and a reproducing apparatus for the same.

To achieve the above object of the present invention, according to one aspect of the present invention, there is provided a method of reproducing audio data by filtering the audio data in response to the fastness or a slowness of a playback speed designated by a user. In the method of reproducing audio data by filtering, a time scale modulation is performed with respect to the audio data in accordance with a predetermined time scale modulation algorithm to increase or decrease the data quantity of the audio data in response to the fastness or slowness of the designated playback speed. Subsequently, either a down-sampling or up-sampling is performed with respect to the audio

data obtained via the time scale modulation in accordance with the fastness or slowness of the designated playback speed to restore the quantity of the audio data after performing the sampling to a level almost the same as the decoded audio data.

Paragraph at page 6, line 23 to page 7, line 5:

In more detail, the sampling step includes the steps of: with respect to the audio data stored in the middle queue, performing the up-sampling processing when the designated playback speed is faster than the normal playback speed, performing the down-sampling when the playback speed is slower than the normal playback speed, wherein the quantity of the sampled audio data to be transferred to an output queue becomes substantially identical with the quantity of the original audio data; and transferring the sampled audio data stored in the output queue to the buffer means in the set unit per predetermined time interval.

Paragraph at page 15, line 22 to page 17, line 1:

The output data obtained from audio decoder 18 after executing the de-compression and decoding is temporarily stored in an output buffer 24 (FIG. 7) in the packet unit. Here, it is supposed that the user designates the playback speed to a low speed reproduction (e.g., slow by two times) or high speed reproduction (e.g., fast by two times). The audio data recorded on output buffer 24 becomes the data (corresponding to FIG. 9(b)) which is modified in time scale to respectively have the modified presentation time interval by responding to the changed playback speed when compared with the data (corresponding to FIG. 9(a)) decoded during the normal playback. For this operation, the MPEG reproducing apparatus carries out a processing for newly setting the presentation time interval by extending or shortening it in response to the fast or slow mode of the playback speed designated by the user. That is, it is necessary to carry out a processing in a manner that a playback speed control ratio α between the playback speed designated by the user and normal playback speed is calculated, and the audio data presentation time interval of the normal playback speed is multiplied by playback speed control ratio α to produce a new audio data presentation time interval. The audio signal reproducing apparatus proposed by the present invention is provided with a means, i.e., a program that newly produces the presentation time interval of respective audio data responding to the fastness or slowness of the designated playback speed whenever the user changes the playback

speed via a key input unit (not shown) of the reproducing apparatus. And, the audio data subjected to the filtering process according to the present invention is reproduced in accordance with the calculated presentation time interval. Thus, the program provided to the reproducing apparatus is executed by a control means such as a CPU (not shown). Here, a value of the playback speed control ratio α becomes 1.5 when the low speed reproduction slower than the normal playback speed by 1.5 times is instructed, or becomes 0.5 when the high speed reproduction faster than the normal playback speed by two times is instructed. In other words, the playback speed control ratio α is determined by a reverse relation of a speed ratio between the designated playback speed and normal playback speed.

Paragraph at page 17, line 6 to page 19, line 3:

The filtering process of the audio data carried out by RTTSM filter 22 is schematically shown in the flowchart of FIG. 3. Functions of RTTSM filter 22 may be embodied in software or hardware. The functions of RTTSM filter 22 will be first described with reference to the flowchart of FIG. 3.

A primary function conducted by RTTSM filter 22 is to increase/decrease the data quantity of the audio data of an input queue Qx provided from output buffer 24 in response to the fast or slow playback speed designated by the user, which is the time scale modification (TSM) of the audio data, and storing it to a middle queue Qy as a TSM signal $y(\cdot)$. The TSM of the audio data may be performed by using one of the known TSM algorithms without any particular modifications or with some modifications for a conformity with a target application.

Several audio signal processing techniques have been suggested for adjusting the playback speed of the audio signal as designated by a user. Particularly, there are some known audio signal processing techniques which are capable of varying the playback speed by increasing or decreasing the data quantity on a time scale basis while maintaining the characteristics similar to those inherent in the original audio signal. Among them, an overlap-addition (OLA) algorithm proposed by Roucus and Wilgus in 1985 may be a representative technique. The OLA algorithm has been developed into the synchronized OLA (SOLA), the waveform similarity based OLA (WSOLA), etc. In addition, the techniques that modify or improve the OLA algorithm such as the global and local search time-scale

modification (GLS-TSM), the time-domain pitch-synchronized OLA (TD-PSOLA) and the pointer interval control OLA (PICOLA) have been known.

The description of the present invention hereinbelow utilizes the WSOLA technique as one of the RTTSM algorithms. In accordance with the WSOLA algorithm, the audio data is cut into many blocks by using a window of a predetermined size so that two successive blocks are overlapped by a regular interval, and then the blocks are added after being rearranged by the intervals corresponding to a speed variation to convert the original signal into the data increased or decreased in time scale. So, the WSOLA algorithm can produce the converted signals capable of being reproduced at a speed different from the original playback speed. However, if the signals of mutually different blocks are simply added after changing the time scale intervals, they will be changed to have a sound quality degraded greatly relative to that of the original signal. For allowing the sound quality of the time scaled modified signal to be maximally similar to that of the original signal, when the blocks are rearranged, it is needed that a correlation enabling to determine a waveform similarity between two signals is estimated while providing a minute adjustment interval within a certain range to a required base interval. Then, two block signals are synthesized by moving them as long as a minute adjustment interval corresponding to a value having the greatest waveform similarity. By doing so, it is possible for the sound quality to maintain a level almost similar to that of the original sound regardless of the varying the playback speed. The WSOLA algorithm is based on the above-described concept. That is, the WSOLA algorithm is characterized in that in order to prevent the degradation of the sound quality in the synthesis of the blocks by the rearranging, signals of the two successive blocks are moved by an interval which allows the waveform similarity between two overlapped portions of the two successive blocks to have a maximum value.

Paragraph at page 19, line 14 to page 19, line 19:

For processing the RTTSM filtering applied with the WSOLA algorithm, first, it is periodically checked whether a user has changed the playback speed (step S10). If there is no instruction of changing the playback speed, the processing is performed in accordance with the already-set playback speed. If there is an instruction of changing the playback speed, the reproducing apparatus produces an event.

Paragraph at page 20, line 19 to page 21, line 1:

After the algorithm executing environment is established to correspond to the new playback speed designated by the user, RTTSM filter 22 increases or decreases the data quantity responding to the designated playback speed by using the WSOLA algorithm with respect to the decoded audio data previously stored in buffer 24 having been processed by audio decoder 18. Then, the data is again down-sampled or up-sampled to be returned to buffer 24. Hence, the data supplied to audio output 20 is the data which have been processed by the WSOLA algorithm with down-sampling (or up-sampling).

Paragraph at page 21, line 6 to page 21, line 23:

The RTTSM filtering processing with respect to respective audio packets is attained by performing three functions which are the RTTSM-put function, RTTSM-calc function and RTTSM-out function. The RTTSM-put function reads out audio data (corresponding to FIG 9(b)) by one set from buffer 24 to write it in input queue Qx (step S18). The RTTSM-calc function performs the WSOLA algorithm processing upon the audio data accumulated on input queue Qx in the frame unit to increase or decrease the data quantity in response to the designated playback speed. So, the time-scaled audio data $y(\cdot)$ (corresponding to FIG. 9(c)) having the increased or decreased data quantity by responding to the current playback speed is formed to be written on middle queue Qy. The audio data accumulated on middle queue Qy is down-sampled for reducing the data quantity again when the currently-designated playback speed is slower than the normal playback speed or is up-sampled for increasing the data quantity when the currently-designated playback speed is faster than the normal playback speed, and the sampled data is written on output queue Qx (step S20). Also, the RTTSM-out function again supplies the audio data accumulated on output queue Qz to buffer 24 by sets, thereby replacing the existing audio data supplied from audio decoder 18 with the data obtained after performing the RTTSM filtering process (step S22).

Paragraph at page 22, line 9 to page 23, line 2:

The audio packet newly obtained by carrying out the RTTSM algorithm is reproduced by audio output 20 to have a tone substantially identical to that of the normal playback, with no

dependency on the playback speed designated by the user. The reason of obtaining such result will be described with reference to FIGS. 4 to 10.

FIG. 9 provides views showing, when the designated playback speed is slower than the normal playback speed by two times, changes to the presentation time interval of the audio data per respective data processing steps. FIG. 9(a) shows the presentation time interval of the audio data corresponding to the normal playback speed. Assuming that the presentation time interval of respective audio data d1, d2,.., d10, ... is t during the normal playback, audio decoder 18 generates the data which has the presentation time interval of respective audio data d1, d2,..., d10,... simply increased by two times as shown in FIG. 9(b) and stores the generated data in buffer 24. Since the presentation time interval of respective audio data d1, d2,..., d10, ... stored in buffer 24 is 2 t, the reproducing time of the audio data is also expanded by two times. If the presentation time interval of the audio data is increased by two times in time scale, the tone of the reproduced sound is lowered roughly by one octave with the consequence of deteriorating the quality of the reproducing sound although the user's desired playback speed can be satisfied.

Paragraph at page 23, line 18 to page 24, line 3:

In order to solve these problems, the audio data obtained after performing the WSOLA algorithm is subjected to the down-sampling. For performing the down-sampling, it is conceptually assumed that the presentation time interval of the audio data is compressed in the time scale to be restored to t as shown in FIG. 9(d) with respect to the audio data obtained after performing the WSOLA algorithm. Once such a processing is carried out, the total reproducing time becomes that as shown in FIG. 9(b). Accordingly, the audio data can be reproduced to conform to the new playback speed set by the user and to be synchronized with the video data. In addition, since there is an effect of recompressing by 1/2 in time scale, the tone of the audio data is raised by one octave to be restored to be almost identical with the tone as shown in FIG. 9(a).

Paragraph at page 24, line 14 to page 25, line 21:

Because the audio data shown in FIG. 9(e) is obtained by down-sampling the audio data (corresponding to FIG. 9(d)) having the tone raised by one octave after compressing the audio data of FIG. 9(c) by half in time scale, the tone thereof is still identical with the tone of the audio data of

FIG. 9(d), which is in turn identical with the tone of the audio data of FIG. 9(a). Consequently, while the playback speed is slowed by two times, the tone of the reproduced sound is maintained to be almost the same as that in the normal playback. Of course, the resolution of the audio data is degraded while performing the down-sampling, but the deterioration of the sound quality caused by the degraded resolution is negligible once a sound quality lowering method to be described later is applied during performing the down-sampling.

FIG. 10 provides views showing, when the designated playback speed is faster than the normal playback speed by two times, changes of the presentation time interval of the audio data per respective data processing steps. FIG. 10(a) shows the presentation time interval of audio data S1, S2,..., S10, ... during performing the normal playback. When the two-fold fast playback is instructed by the user, the reproducing apparatus compresses the sample presentation time interval of respective audio data by 1/2, i.e., $t \rightarrow t/2$, as shown in FIG. 10(b). The audio data stored in buffer 24 is to be reproduced by the time interval of t/2 when being reproduced as it is. Accordingly, the tone of the reproduced sound is to be raised by one octave as compared with that of the normal playback. Therefore, the audio data is processed in such a manner that the WSOLA processing and up-sampling are executed with respect to the data stored in buffer 24 to not only quicken the playback speed by two fold but also maintain the tone of the normal playback in the reproduced sound.

Firstly, the data stored in buffer 24 is subjected to the WSOLA processing to decrease the quantity of the audio data by substantially 1/2 as shown in FIG. 9(c). At this time, since the presentation time interval of respective audio data continuously maintains t/2 unchanged, the tone also maintains the state of being raised by one octave as compared with that of the normal playback. The reproducing time of the audio data after performing the WSOLA processing is shortened by as much as 1/4 as compared with that of the normal playback causing the problem of inconsistent synchronization with the video data as well as the problem of maintaining the tone variation higher by one octave.

Paragraph at page 26, line 8 to page 27, line 19:

However, the number of audio data samples is still only one-half that shown in FIG. 10(b), and the reproducing apparatus is prearranged to present the audio data per t/2. Due to these facts,

only the compression in time scale is insufficient. In other words, for reproducing the audio data in accordance with the presentation time interval of t/2, it is required for the audio data obtained by performing the WSOLA processing shown in FIG. 10(c) to have the quantity increased by two times. For this purpose, the up-sampling is performed with respect to the audio data obtained from the WSOLA processing, so that its data quantity is increased by two times. By performing the up-sampling, the audio data as shown in FIG. 10(e) is finally obtained.

Because the audio data S1″, S2″, ..., S10″... shown in FIG. 10(e) is obtained by up-sampling upon the audio data (corresponding to FIG. 10(d)) having the tone lowered by one octave after expanding the audio data of FIG. 10(c) by two times in time scale, the tone thereof is still identical with the tone of the audio data of FIG. 10(d), which is in turn identical with the tone of the audio data of FIG. 10(a). Consequently, while the playback speed is quickened by two times, the tone of the reproduced sound is maintained to be almost the same as that of the normal playback.

The above-described down-sampling or up-sampling after executing the WSOLA algorithm is performed by three functions which will be described later. Also, the down-sampling or up-sampling is performed in a manner that the increase or decrease rate of the data is determined in accordance with the fastness or slowness of the playback speed designated by the user, and the quantity of the audio data is increased or decreased in accordance with the determined increase/decrease rate. Amplitudes of the respective audio data after the sampling may take those of the TSM audio data obtained from the WSOLA processing unchanged or may be determined by interpolating the amplitudes of the adjacent audio data. Herein below, a specific data processing algorithm by using respective functions will be described.

FIGS. 4, 5 and 6 are flowcharts respectively showing the routines of the RTTSM-put function, RTTSM-out function and RTTSM-calc function, and FIG. 7 is a view illustrating a process of transforming respective audio packets of buffer 24 into new audio packets via input queue Qx, middle queue Qy and output queue Qz by implementing the three functions. FIG. 8 provides views illustrating a principle of obtaining a TSM signal y(×) such that the length of original audio signal x(×), i.e., the quantity of the audio data, is expanded or compressed in time scale in response to the fastness or slowness of the playback speed set by the user. In the present invention, three queues are utilized for performing the WSOLA processing and the up/down-sampling using the three functions.

Paragraph at page 28, line 7 to page 29, line 16:

Input queue Qx is preferably required to have a size long enough for accumulating the audio data of more than roughly 3 frames. As one set is written, a pointer value of input queue Qx is increased. After the queue pointer indicates the last position of input queue Qx during the process of increasing the queue pointer, it is reset to indicate the starting position to allow input queue Qx to serve as a circular queue. In addition, as one set is written on input queue Qx, it is counted. Then, as the counted number of sets becomes the same as the set value of parameter S_a, a calc-nextframe flag for deciding whether the next frame is calculated or not is changed to Enable. Of course, the default value of the calc-nextframe flag is set as Disable, and the change of the value to Enable denotes that input queue Qx is stored with at least one frame capable of performing the WSOLA algorithm.

Together with writing the audio data before performing the filtering according to the present invention on input queue Qx by reading out from buffer 24 by one sets, RTTSM-out function as shown in FIG. 5 is carried out to read out the audio data stored on output queue Qx having been subjected to the WSOLA processing and up/down-sampling processing by one sets d_{ij} and then overwrite it on buffer 24 in the same rate of the input case as the set index is increased by one (step S36). Because the data quantity after performing the WSOLA processing and down/up-sampling processing is the same as that prior to performing the processings, no problem occurs except for the postponing of the overall reproducing time for a short time period (i.e., time required for performing the WSOLA processing and down/up-sampling processing) even though the data is read out in sets from output queue Qz to be sequentially written on buffer 24. Output queue Qz is set to have a size capable of being simultaneously stored with the data of at least two frames, and the queue pointer is adjusted for serving as the circular queue (step S38).

During transmitting the audio data accumulated on input queue Qx to output queue Qx, the RTTSM-calc function as shown in FIG. 6 is executed to perform the TSM processing based on the WSOLA algorithm and down/up-sampling processing. It should be noted that, while the execution period of RTTSM-put function and RTTSM-out function is of the set unit, the execution period of the RTTSM-calc is processed in the frame unit which is a group of a plurality of sets. That is, the RTTSM-calc function is implemented only when the value of calc-nextframe flag is in the Enable state (step S40). Also, whenever the foregoing processing upon the current frame is carried out, the

value of calc-nextframe flag is shifted to Disable to prepare the processing of the next frame (step S42).

Paragraph at page 30, line 5 to page 30, line 11:

When there is no change in the playback speed, the WSOLA processing is performed with the preset values of environmental parameters as follows. By executing the RTTSM-put function, the input queue Qx is accumulated with the audio data. Here, the RTTSM processing with respect to the audio data stored in input queue Qx is performed every time the calc-nextframe flag is set to Enable. In order to perform the WSOLA processing, it is required for input queue Qx to be stored with audio data of at least one frame.

Paragraph at page 32, line 14 to page 32, line 22:

Here, in computing the optimum correlation between successive frames, a computing method with sliding the audio data one by one is available. However, this computing method imposes a burden of performing a lot of calculations on the reproducing system. Therefore, a method of skipping a plurality of audio data may be recommendable as the computing method of the optimum correlation when it is required to speed up the calculating speed. However, it is inevitable that the method would be inferior to the former method in view of an accuracy of the optimum correlation. It is preferable to consider a performance of a CPU of the reproducing apparatus in deciding which method would be more suitable.

Paragraph at page 34, line 2 to page 34, line 4:

Here, g(j) is a weighted value function, of which a representative form is preferably a linear function. Alternatively, an exponent function may also be applied as the weighted value function.

Paragraph at page 34, line 16 to page 35, line 6:

The audio data accumulated in middle queue Qy via the WSOLA processing is then transferred to output queue Qz. During the transferring, the down-sampling or up-sampling is performed in accordance with the playback speed. In performing the sampling, a data increase/decrease rate is determined based on the playback speed designated by the user, and then

the audio data quantity is varied in accordance with the determined increase/decrease rate by using an interpolation method capable of not causing any changes in data characteristics before and after the sampling. The interpolation method is a numerical analysis method for inferring a new point from other given points. There are some typical interpolation methods: the interpolation method using the Taylor polynomial which is commonly employed in numerical interpretation, the interpolation method using the Lagrange polynomial, the repetitive interpolation method, the Hermite interpolation method and the three-dimensional Spline interpolation method, and a linear interpolation method which is the simplest one. Any interpolation method may be applied to the present invention only if it allows the characteristics of the audio data to be almost identical to each other before and after the sampling.

Paragraph at page 38, line 13 to page 37, line 14:

It is worthwhile to generalize the above method to be modified and applied to the case where the playback speed control ratio α has any other values.

Paragraph at page 39, line 12 to page 39, line 18:

The audio data newly obtained by the down-sampling or the up-sampling is transferred to output queue Qz in the frame unit. And the audio data of the output queue Qz is sequentially written to buffer 24 by sets by the execution of the RTTSM-out function. By doing so, an existing audio packet of buffer 24 is replaced with a new corresponding audio packet from output queue Qz that has been subjected to the WSOLA processing and down/up-sampling. The audio data to be provided to audio output 20 is the new corresponding audio packet.

Paragraph at page 40, line 3 to page 41, line 9:

The present invention introduces three data storage means which are input queue, middle queue and output queue for the TSM processing and up/down-sampling processing. But it should be appreciated that there is no need to separate them in the physical sense as one memory of the reproducing apparatus may be divided into three memory areas and so utilized. Furthermore, three queues are defined for the convenience of embodying the software but there is no need to define three queues separated as above. In other words, there may be other ways of defining the queues that

form one unified full-size queue of which is divided into three and each of the three regions is defined to act as a circular queue by controlling a pointer thereof.

The method of processing the audio data according to the present invention as described above can be embodied in a software method to be directly applied to a computer which is installed with the Windows operating system and a program referred to as the Direct Media of Microsoft Co. Ltd. In realizing the software method, the program embodying the algorithm of the audio data processing method is stored in the hard disc (not shown) or a ROM 240 within the computer and is implemented by CPU 230 when a multimedia reproducing program is run. Buffer 24 or three circular queues Qx, Qy and Qz appropriately utilize the resources of a RAM (not shown) within the computer, and a sound card (not shown) within the computer is utilized as the audio output 20.

The possibility of applying the method of processing the audio data according to the present invention is not limited to a computer. The method can be also applied to DVD system 100a, digital VCR system or another similar systems, i.e., any digital reproducing apparatus for reproducing the compressed and encoded video data and audio data. Moreover, it may be applied to a tape recorder, VCR system 100b of analog system, or similar system. In other words, the method of processing the audio data according to the present invention can be widely applied regardless of the analog system or digital system without being related to the compressing method or encoding method of data once it is for a reproducing apparatus related to the processing of audio data. Just that, in terms of the reproducing apparatus of analog system, the audio signal is converted into a digital signal, the RTTSM filtering processing according to the present invention is performed, and it is converted to the analog signal again to be reproduced.

Paragraph at page 41, line 19 to page 43, line 14:

Naturally, the reproducing apparatus is provided for the purposes of the present invention with a playback speed control means for calculating the playback speed control ratio α between the user's designated playback speed and the normal playback speed and calculating the new presentation time interval after multiplying the audio data presentation time interval of the normal playback mode by playback speed control ratio α. A combination of a key input (not shown) and a controller such as a microcomputer and a CPU 230 can function as the playback speed control means.

DSP board 200 may consist of a ROM 240, a RAM (not shown) in which three queues can be formed by defining the RAM resource, CPU 230 or DSP chip, an oscillator (not shown), an analog/digital converter (ADC) 210, a digital/analog converter (DAC) 220, and so on. A program realizing the RTTSM-calc function is resident in ROM 240, and the RAM is operated to be utilized as input queue Qx', middle queue Qy' and output queue Qz'. ADC 210 is supplied with audio signals recorded on the video tape from a servo 100 to convert it into digital data. DAC 220 converts the digital data into analog signals to permit it to be reproduced as sound via speaker 300. CPU 230 sequentially implements the loaded program stored in ROM 240 to perform several data processing tasks for writing the output data of ADC 210 on input queue Qx', transferring audio data accumulated in output queue Qz' to DAC 220 and performing the WSOLA processing and the down/up-sampling with respect to audio data by implementing the above-stated RTTSM-calc function with respect to the data accumulated on input queue Qx'. When the source signal recorded on the recording medium is recorded as the analog signal, as in the analog VCR, ADC 210 is necessary. But, ADC 210 is not required when the source data is of the digital signal as in the DVD system.

DSP board 200 is formed with a background 200a and a foreground 200b. Background 200a performs the functions of processing the audio data on the hardware basis, writing the output data of ADC 210 on input queue Qx' and transmitting the audio data accumulated on output queue Qz' to DAC 220. The foreground 200b performs the function of transferring the data obtained by performing the WSOLA processing and the down/up-sampling in turn with respect to the audio data stored in input queue Qx' by implementing the RTTSM-calc function in accordance with the program to the output queue Qz'. That is, background 200a plays the roles of foregoing RTTSM-put function and RTTSM-out function on the hardware basis. In other words, background 200a simultaneously performs a writing operation of the audio data of an audio signal supplying means 100a or 100b to input queue Qx' in the set unit and a reading operation of the audio data stored in output queue Qz' in the set unit, and converts the audio data read out from output queue Qz' as the analog signal. Foreground 200b serves for performing the TSM processing by using a predetermined TSM algorithm like WSOLA with respect to the audio data stored in input queue Qx' in the frame unit to increase/decrease the data quantity in response to the fastness or slowness of the designated playback speed, and performing the down-sampling or up-sampling with respect to the audio data

obtained via the TSM processing in accordance with the designated playback speed to restore the quantity of the audio data after being subjected to the sampling to the level substantially identical with that of the original audio data to transmit it to output queue Qz'.

Paragraph at page 44, line 11 to page 45, line 13:

When the interrupt signal is generated periodically by counting the clock signal provided by an oscillator of the reproducing apparatus, a value of the ISR having the default value as Disable is shifted into Enable, and data processing (steps S64 to S72) by background 200a is carried out whenever the ISR is Enabled. Because foreground 200b performs the filtering processing upon the audio data obtained by carrying out the ISR of background 200a, an infinite loop is implemented until a next-frame-start flag is shifted into Enable (step S74).

In order to perform the ISR processing, CPU 230 brings out the audio data of one set from ADC 210 (step S64), and separately brings out a playback speed designated by a user from the user interface such as the key input (not shown). The audio data from ADC 210 is written on input queue Qx' (step S66). A value is cumulatively counted as writing it on input queue Qx' by one set at a time, and it is checked whether the counted value reaches the total set number included in a single frame. If it is true, a value of the next-frame-start flag, which is initially set to Disable, is shifted into Enable (steps S68 and S70). The processing hereinbefore is equivalent to that of the above-stated RTTSM-put function. The difference is that the output data of ADC 210 is written on input queue Qx'. Subsequently, CPU 230 accesses the output queue Qz' to read out one set of the audio data stored therein to transfer it to DAC 220 (step S72). This is equivalent to the RTTSM-out function. The ISR processing as described above is performed only when a background pulse maintains a high state as shown in FIG. 15(b).

The foreground processing is designed to implement an infinite loop once it is initiated. In more detail, if the value of next-frame-start flag is set to Enable, the value of next-frame-start flag is shifted to Disable which is the basic set value (step S76). Thereafter, the RTTSM-calc function is executed upon the audio data stored in input queue Qx' in accordance with the foregoing method to perform the WSOLA processing and down/up-sampling (step S78). Then, the processed data is transferred to output queue Qz' and stays therein until it is outputted to DAC 220.

Paragraph at page 46, line 3 to page 47, line 11:

On the other hand, when being applied to the digital VCR system, overall data processing system is almost the same as the foregoing case except for the slight difference that ADC 210 is not needed in DSP board 200 since the original signal is digital. Similarly, DSP board 200 may be formed without employing ADC 210 due to the fact that this original signal is the digital signal regardless of a difference that the recording medium of the DVD system is the DVD without being the tape, and the overall data processing is almost the same as in the foregoing case.

According to one aspect of the present invention as described hereinbefore, the audio data is reproduced by applying the method of extending/compressing the value of the presentation time interval of respective audio data in accordance with a value of the designated playback speed. According to the above method, since the audio data should be reproduced and output by corresponding to the designated presentation time interval, the process of down-sampling or up-sampling upon the audio data is required.

However, according to another aspect of the present invention, audio output 20 is controlled to extend/compress a whole presentation time of the audio data in accordance with the fastness or slowness of the designated playback speed while maintaining the presentation time interval of respective audio data as the value of the normal playback speed. According to this aspect, the down-sampling or the up-sampling is not required in case of the slow playback mode or the fast playback mode. More specifically, it is controlled so that the whole presentation time of the audio data set by the normal playback speed as a reference is extended/compressed in response to a value of the designated playback speed, and the presentation time interval of the audio data maintains the value of the normal playback speed. Meanwhile, the TSM processing is performed with respect to the audio data by applying the above-described TSM algorithm to increase/decrease the data quantity in accordance with a value of the playback speed designated by the user. Then, the audio data subjected to the TSM is controlled to be reproduced during the changed presentation time by the presentation time interval. Once the signal processing for reproducing the audio data is performed in the foregoing manner as described above, the reproduced sound also maintains the tone substantially identical with that of the normal playback speed without being influenced by the value of the designated playback speed. It is advantageous in that the sampling of the audio data can be deleted to allow the sound quality to be nearer to the original sound.

Paragraph at page 47, line 21 to page 48, line 3:

Furthermore, the method of processing the audio data according to the present invention may be performed independently of the processing of the video data. Therefore, it is widely applicable to above-mentioned, different media reproducing apparatuses. In other words, a module embodied with the method of processing the filtering of the audio data according to the present invention is simply added to an audio signal processing module of respective media reproducing apparatuses, thereby being capable of forming the media reproducing apparatus to have the audio data reproducing function according to the present invention.

CLAIMS (with indication of amended or new):

---

1. (AMENDED) A method of reproducing original audio data having a given sampling quantity and a given tone, in response to a value of a playback speed designated by a user, comprising the steps of:

performing a time scale modulation processing with respect to the original audio data in accordance with a time scale modulation algorithm to increase or decrease the quantity of the original audio data in response to the value of the playback speed; and

down-sampling or up-sampling with respect to audio data obtained by the time scale modulation processing in accordance with the value of the designated playback speed to restore the quantity of sampled audio data to a level of the given sampling quantity of the original audio data in a manner such that a tone of the sampled data is substantially identical to the given tone of the original audio data while the sampled data is reproduced at the playback speed designated by the user.

2. (AMENDED) A method of reproducing audio data as claimed in claim 1, further comprising newly calculating a presentation time interval of the audio data to be increased/decreased in accordance with the value of the designated playback speed in response to a change of the playback speed.

3. (AMENDED) A method of reproducing audio data as claimed in claim 2, further comprising reproducing the sampled audio data by a newly-calculated presentation time interval.

4. (AMENDED) A method of reproducing audio data as claimed in claim 1, wherein the step of time scale modulation comprises the steps of:

writing the original audio data stored in a buffer on an input queue in a set unit per predetermined time interval; and

a1 cont'd

performing the time scale modulation algorithm in a frame unit upon the audio data stored in the input queue to decrease the quantity of the audio data in accordance with the designated playback speed when the designated playback speed is faster than the normal playback speed, or to increase the quantity of the audio data in accordance with the designated playback speed when the designated playback speed is slower than the normal playback speed, and providing time scaled audio data to a middle queue.

5. (AMENDED) A method of reproducing audio data as claimed in claim 4, wherein the sampling step comprises the steps of:

with respect to the time scaled audio data stored in the middle queue, performing the up-sampling processing when the designated playback speed is faster than the normal playback speed, performing the down-sampling when the playback speed is slower than the normal playback speed, so that the quantity of the sampled audio data to be transferred to an output queue is substantially identical to the given sampling quantity of the original audio data; and

transferring the sampled audio data stored in the output queue to the buffer in the set unit per predetermined time interval.

a2

7. (AMENDED) A method of reproducing audio data as claimed in claim 5, wherein the sampled audio data of the output queue is overwritten to the buffer so as to replace the original audio data existing in the buffer.

a3 cont

9. (AMENDED) A method of reproducing audio data as claimed in claim 4, wherein the number of sets of the original audio signal which is written to the input queue is cumulatively

a3 cont'd

counted, and a calc-nextframe flag having a Disable default state is shifted to an Enable state when the counted number of sets becomes equal to the number of sets of one frame, thereby performing the time scale modulation algorithm in the frame unit.

a4 cont

11. (AMENDED) A method of reproducing audio data as claimed in claim 1, wherein in the up/down sampling, a varying ratio of data quantity is calculated in accordance with the value of the designated playback speed, and the quantity of the audio data obtained by the time scale modulation processing is varied in accordance with the varying ratio while characteristics of the audio data before and after the up/down-sampling are substantially identically maintained by using data interpolation.

12. (AMENDED) A method of reproducing audio data as claimed in claim 1, wherein the time scale modulation algorithm increases or decreases the quantity of the original audio data in accordance with the value of the designated playback speed while maintaining characteristics of the original audio data.

13. (AMENDED) A method of reproducing decoded audio data in response to a playback speed designated by a user, before supplying the decoded audio data, which has been stored in a storage and been decoded in the MPEG system, to an audio output, comprising the steps of:

calculating a playback speed control ratio between the designated playback speed and a normal playback speed, and multiplying a presentation time interval of the decoded audio data in case of the normal playback speed by the playback speed control ratio to produce a new presentation time interval of the audio data;

writing the decoded audio data stored in the storage on an input queue in set units;

performing a time scale modulation algorithm in a frame unit with respect to audio data written on the input queue to increase or decrease a quantity of the decoded audio data in proportion to the playback speed control ratio, where audio data after the time scale modulation processing is written on a middle queue;

with respect to the audio data written in the middle queue, performing an up-sampling in case of a fast playback mode where the playback speed control ratio is smaller than 1 or a down-sampling

in case of a slow playback mode where the playback control ratio is larger than 1, in a manner such that a sampling rate is applied as a reverse number of the playback speed control ratio for allowing the quantity of the audio data after performing the sampling to be substantially identical to the decoded audio data and sampled audio data is transferred to an output queue;

writing the audio data stored in the output queue to the storage in the set unit to replace existing decoded audio data; and

reproducing the audio data newly written to the storage by the produced presentation time interval, such that a tone of a reproduced sound is substantially identical with that of the normal playback speed even when the designated playback speed is faster or slower than the normal playback speed.

---

15. (AMENDED) A method of reproducing audio data as claimed in claim 12, wherein the set unit is comprised of one audio data in case of a mono system or of two audio data for left/right channels in case of a stereo system.

---

17. (AMENDED) A method of reproducing audio data as claimed in claim 12, wherein the time scale modulation algorithm increases or decreases the quantity of the decoded audio data in accordance with a value of the designated playback speed while maintaining audio characteristics of the decoded audio data.

18. (AMENDED) A method of reproducing audio data after being subjected to a filtering processing in accordance with a value of a playback speed designated by a user, comprising the steps of:

increasing or decreasing a presentation time of the audio data having a normal playback speed in response to the value of the designated playback speed, and maintaining a presentation time interval of the audio data to have a value of the normal playback speed;

performing a time scale modulation processing by using a predetermined time scale modulation algorithm with respect to the audio data to increase or decrease a quantity of the audio data in accordance with the value of the designated playback speed; and

reproducing the audio data obtained from the time scale modulation processing during the changed presentation time by the presentation time interval, such that a tone of a reproduced sound is substantially identical to that of the normal playback speed even when the designated playback speed is faster or slower than the normal playback speed.

19. (AMENDED) A method of reproducing audio data as claimed in claim 18, wherein the predetermined time scale modulation algorithm increases or decreases the quantity of the decoded audio data in accordance with the value of the designated playback speed while maintaining audio characteristics of the decoded audio data.

20. (AMENDED) An apparatus for reproducing audio data in response to a value of a playback speed designated by a user, comprising:

a playback speed control that produces a playback speed control ratio between the designated playback speed and a normal playback speed, and a new presentation time interval by multiplying a presentation time interval of the audio data at the normal playback speed by the playback speed control ratio;

a storage for storing the audio data in packet units;

a filter that provides time scale modulation processing in accordance with a predetermined time scale modulation algorithm with respect to the audio data stored in the storage to increase or decrease a data quantity of the audio data in accordance with the value of the designated playback speed, the filter further provides a down-sampling or up-sampling with respect to audio data obtained from the time scale modulation processing in accordance with the value of the designated playback speed to restore the quantity of sampled audio data to a level substantially identical with that of the audio data prior to the time scale modulation processing, and the filter writes the sampled audio data to the storage to replace existing audio data; and

an audio output which receives the filtered audio data from the storage by a new presentation time interval and reproduces the filtered audio data into a sound, such that a tone of a reproduced sound is substantially identical with that of the normal playback speed even when the designated playback speed is faster or slower than the normal playback speed regardless of being reproduced by the new presentation time interval.

21. (AMENDED) An apparatus for reproducing audio signals as claimed in claim 20, wherein the predetermined time scale modulation algorithm increases or decreases the quantity of the audio data in accordance with the value of the designated playback speed while maintaining audio characteristics of the audio data.

22. (AMENDED) An apparatus for reproducing audio signals as claimed in claim 20, wherein in the up/down sampling, the filter calculates a varying ratio of data quantity in accordance with the value of the designated playback speed, and varies the quantity of the audio data obtained by the time scale modulation processing in accordance with the varying ratio while substantially identically maintaining audio characteristics of the audio data before and after the up/down sampling by using data interpolation.

23. (AMENDED) An apparatus for reproducing audio signals comprising:

an audio signal supplier that provides audio signals from a recording medium in response to a value of a playback speed designated by a user; and

a digital signal processor having a background portion for simultaneously writing audio data of the audio signal supplier on an input queue in set units and reading out of the audio data stored in an output queue in a set unit referenced to a frame unit, and converting the audio data read out from the output queue into an analog signal, and a foreground portion for performing a predetermined time scale modulation by using a predetermined time scale modulation algorithm in the frame unit with respect to the audio data stored in the input queue to increase or decrease the data quantity in response to the value of the designated playback, performing a down-sampling or up-sampling with respect to the audio data obtained by the time scale modulation processing in accordance with the value of the designated playback speed to restore a quantity of the sampled audio data to a level substantially identical with that of the audio data prior to the time scale modulation, and transferring the sampled audio data to the output queue.

24. (AMENDED) An apparatus for reproducing audio signals as claimed in claim 23, wherein the digital signal processor further comprises an analog/digital converter for converting an analog

audio signal into digital data between the audio signal supplier and the input queue when the audio signal supplied from the audio signal processor is an analog signal.

25. (AMENDED) An apparatus for reproducing audio signals as claimed in claim 23, wherein the predetermined time scale modulation algorithm increases or decreases the quantity of the audio data in accordance with the value of the designated playback speed while maintaining audio characteristics of the audio data.

26. (AMENDED) An apparatus of reproducing audio signals as claimed in claim 23, wherein in the up/down sampling, the digital signal processor calculates a varying ratio of data quantity in accordance with the value of the designated playback speed, and varies the quantity of the audio data obtained by the time scale modulation processing in accordance with the varying ratio while substantially identically maintaining audio characteristics of the audio data before and after the up/down sampling by using data interpolation.